David W. Blois, Ph.D.
Senior Vice President
Global Regulatory Policy

Merck & Co., Inc.
West Point PA 19486
E-Mail: david_blois@merck.com
Tel 484 344 2304
215 652 5000
Fax 484 344 2335

May 28, 2003

Dockets Management Branch (HFA-305)
Food and Drug Administration
5630 Fishers Lane, Rm. 1061
Rockville, MD 20852

RE: [Docket No. 02N-0528 - Risk Management - Request for Comments]

Merck & Co., Inc, is a leading worldwide, human health product company. Merck Research Laboratories (MRL), Merck's research division, is one of the leading U.S. biomedical research organizations. Merck's R & D pipeline has produced many of the important pharmaceutical products on the market today.

Merck supports regulatory oversight of product development and welcomes guidance for compliance that is based on sound scientific principles and good judgment. As a leading pharmaceutical company, Merck has extensive experience in thoroughly evaluating our products from discovery to approval and throughout their marketing life to assure that they continue to provide health benefits with minimum risk. Therefore, we are well qualified to comment on the risk assessment and risk management concept papers issued by FDA on March 7, 2003[1], entitled *Premarketing Risk Assessment; Risk Management Programs and Planning;* and *Risk Assessment of Observational Data -- Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment.*

General Comment

We commend the FDA for its efforts in the development of guidance for industry on good risk assessment, risk management, and pharmacovigilance practices, and particularly for its issuance of the three concept papers that outline its current thoughts to encourage discussion of this important topic. Because risk assessment and management is a continuum from discovery through the marketing life of a product, and practices adopted in the pre-marketing phase are likely to influence plans for the future, it may be more appropriate for the Agency to consider combining the concepts separately discussed in these concept papers in a single cohesive guidance.

I. Concept Paper I. "Premarketing Risk Assessment"

A. General

1 The concept paper is written with a focus on the ideal. The guidance document will need

to provide advice on the best *attainable* practices and provide context for their appropriate consideration. An inventory of the ideal would be of little value as regulatory guidance.

2. The practicality and, indeed, the public health benefit of certain recommendations are neither apparent nor addressed. According to the concept paper, ideally, all programs would include long-term, controlled safety studies conducted pre-approval and greater heterogeneity in the safety data base, particularly in phase 3 studies, including use of a range of doses. It must be appreciated that these measures would greatly increase study size if expected to provide a meaningful increment in safety information. This would lead to longer pre-market development times, ultimately delaying availability of new therapies to the public and increasing their cost. The current process for pre-market evaluation of new drugs, which may include some or all of these measures when deemed necessary, often in consultation with FDA, is highly successful in bringing safe and effective drugs to market while focusing resources on products that require special consideration for safe use.

3. The concept paper lacks discussion of the advantages and limitations of the practices it discusses or of situations in which such practices should be considered. To be useful, the draft guidance should include such discussion. It should not adopt a "one size fits all" philosophy.

4. There are different risks and benefits to various pre-market risk assessment practices themselves (marginal rate of return of meaningful safety data versus delays in availability and increases in cost of development). When drafting the guidance, FDA needs to avoid an overly risk-averse position and focus on achieving a balance between risk and benefit in its recommended approach. Only this approach will achieve a positive public health gain.

5. The paper appears to have been written without reference to the content of the other concept papers on risk management or pharmacovigilance. A risk management program (RMP) will likely span the pre-market and post marketing time frames. Taken separately, it is difficult to get a clear understanding of how FDA believes activities in these phases complement or, indeed, substitute for each other.

6. The paper does not address assessment of adverse events for intensity, duration, or frequency, all of which are important characteristics of the risk profile that will help to guide the risk management program that is required.

B. Specific Comments

1. II-A. What is risk assessment?

Lines 23-27: FDA states that the process of identifying a product's underlying risks prior to approval "entails **ensuring** that the body of evidence generated by the clinical trials not only defines the product's effectiveness but also **comprehensively** describes its safety (as required by the Food, Drug, and Cosmetic Act, which calls for the conduct of all tests reasonably applicable to evaluate a drug's safety)."

"Ensuring" and "comprehensively" impute an unrealistic expectation with respect to pre-market assessment. It is practically impossible to "ensure" that pre-marketing studies can define the risk of any product to this degree. Indeed, the sentence beginning on line 20 makes the point that risk assessment occurs throughout a product's life cycle.

Recommendation: A more realistic expression of the process for identifying a product's underlying risks might be, "This process entails the conduct of preclinical, non-human studies and clinical trials prior to approval that not only define the product's effectiveness but also provide a thorough assessment of its safety profile."

2. III-A. What is the appropriate size of the premarketing safety database?

Lines 63-136: The ICH[2] published guidance on the size of the pre-marketing safety database for products used chronically in non-life-threatening conditions. FDA was a participating member in the development of that recommendation. This concept paper asks for input on general guidance that could be provided on the appropriate size of safety database for subsets of products (products only for acute use, products only for serious or life-threatening conditions). It also lists reasons, in addition to those discussed in the ICH guidance, why a larger database might be appropriate.

[2] Guideline for Industry: E1A The Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-term Treatment of Non-Life-Threatening Conditions

Lines 79-83 Appropriate database size[3].

Recommendations:

a. The ICH E1A guidance establishes a population exposure benchmark for assessing clinical safety for drugs intended for long-term treatment of non-life-threatening conditions, including a discussion of factors that may call for a larger data base for such products. This is a category of drugs that, because of the need for chronic use for non-life-threatening indications, demands evidence of a reasonably benign safety profile. The stated philosophy of ICH E1 is that the safety evaluation during clinical drug development will characterize well those events with an incidence of about 1% while not expecting to characterize rare adverse events (those occurring in less than 1 in 1000 patients). In addition, ICH E1A is clear that the actual patient exposure for a specific drug will be determined by the information available on the drug, the drug class, and any post-marketing surveillance requirement. We believe the underlying philosophy of ICH E1A is sound and, given the ICH benchmark for drugs indicated for chronic use in non-life-threatening conditions, is applicable to other product categories in making reasonable determinations about population exposure to assess clinical safety. We recommend against attempting to set standard numbers for pre-market exposure in regulatory guidance for therapies of all varying durations and degrees of disease severity (acute use, serious or life-threatening diseases, and others).

[3] The appropriate database size depends on how sensitive the detection process is to be. Even to have only a 50% chance of having at least one event occur when the probability of an event is, say, 1/1000, there must be at least 700 patients at risk; this number increases to 2600 if at least 3 events must occur to provide some assurance that the occurrence is unlikely to be by chance alone, and to 4800 if at least 5 events must occur. The numbers rise quite dramatically as the event rate decreases. Using the ICH estimate of 1500 patients, there is a 95% chance of seeing at least one event if the event rate is 0.2% or more, and a 50% chance if the rate is 0.043% or more. Clearly, unless the event occurs with some appreciable probability, it is unlikely to be seen at all, let alone frequently enough to allow for treatment-control comparisons, unless the sample size is large. However, large sample sizes for important demographic subgroups are not likely to be available premarketing, so it would seem unrealistic to expect that evaluation of risks specific to important subgroups is likely to be informative unless the probability of an adverse event is remarkably large in that subgroup.

The insensitivity becomes worse when between-group comparisons are needed. Very large sample sizes indeed would be required to have sufficient events to be able to conclude that there was a differential treatment effect if the overall event rate were small (less than 1%, say).

This suggests that while (p. 6, lines 167-168) including a more diverse group of patients may help to generalize the findings of the pre-merket development program, it is unlikely to be informative about differential risks among different subpopulations of patients.

Alternatively, one might ask how large a sample is required to have a good chance of making a statement like "There is 95% confidence that the true event rate is no more than 1%". Statements of this sort quantify the upper limit of the public health risk that might be anticipated. If the true event rate really is 0.5%, a population of at least 2000 patients will be needed to have a 90% chance of being able to make the desired confidence statement. As the true event rate gets closer to 1%, the required population size increases dramatically: a population of 20,000 patients is needed to have this much of a chance of making the statement if the true event rate is 0.8%. With rare events, the sample sizes required increase by an order of magnitude. Population sizes of this magnitude are unlikely to be provided by any development program even on an overall basis, let alone for any important subgroups. Elaborate pre-market risk assessment plans therefore are unlikely to be very effective.

b. We do not believe there is a need to add to the list of reasons included in ICH E1A for a larger database. The additional reasons suggested in the concept paper (lines 131-136) are vague and add little to the discussion in the existing guidance. Vaccines, for example, represent a unique product class with an excellent clinical safety record that confirms that current practices for premarket evaluation are highly effective. Therefore, we believe that vaccines represent an example that is out of context in this discussion. The suggestion that a larger database may be necessary when a very safe alternative to the investigational product is available is to suggest that the new product must not only be "safe and effective", but must be as safe and effective or more so compared to existing therapy. Such comparative safety and efficacy is not a requirement under the FD&C Act, and the sample size requirements to provide meaningful data when between group comparisons are needed make such evaluations impractical except in specific situations where it will, no doubt, be recognized without additional formal guidance. In addition, the context of the disease being treated, the pre-clinical and early phase clinical safety assessment of the new drug, the efficacy of the alternative marketed product, and the potential efficacy advantages of the new drug to existing therapy all warrant consideration in determining the appropriate pre-market exposure for a given situation and, in some cases, may obviate the need for a larger population.

c. As discussed in the ICH E1A guidance, the recommendation of 1500 patients is a general benchmark but the actual numbers for a specific development program need to be justified on a case by case basis. It should be most applicable to products with new mechanisms of action. Patient years and duration of exposure should be addressed. A substantially smaller database should be acceptable for compounds in the same class of approved products, or for combinations of products already in the marketplace.

3. III-B (Line 138): "What are some characteristics of an ideal safety database?"

As noted in our "General Comment 1" above, the focus on the ideal is not useful in a guidance document without practical recommendations. The nature of guidance is to take exigencies into account and offer a recommended path forward towards best attainable under the given circumstances.

Recommendation: Guidance, when published should offer practical recommendations for risk assessment during pre-market development, and recognize that there will always be situations that must be considered on a case by case basis.

4. III-B(1): (Lines 143-157): Long-term controlled safety studies

This paragraph takes the general position that it would be preferable, "in most cases" to have long-term controlled safety data (lines 146-149). However, no justification for this conclusion is provided. The record of the industry and of the FDA in developing and approving new products with appropriate labeling confirms that, for the overwhelming majority of new drugs, current practices work very well. Indeed, the paragraph itself notes that "Generally, events that occur rarely and spontaneously...do not need a control group to be interpreted" (lines 152-153). Further, long-term controlled safety studies are not without their own concerns. It should be acknowledged that long-term placebo studies may not be ethical depending on the disease, and active-control studies may not provide a useful

assessment of those adverse events that are shared across compounds. Even as an "ideal," this suggestion is flawed in that the ideal would surely involve consideration of reasonable costs of development, existence of prior safety signals, the likelihood of such studies to provide meaningful additional safety information, and the public health cost of delaying access to new safe and effective therapies.

Recommendation: The guidance document should discuss the benefits and disadvantages of long-term controlled safety studies and describe situations in which they may be recommended in the pre-marketing phase of development in spite of their limitations. Discussion should include the objectives best achieved with such studies and include discussion of power and typical duration. The guidance should not recommend or imply that such studies should be considered for all programs.

5. III-B(3): Development of safety (and effectiveness) data over a range of doses (and plasma levels) throughout the clinical program (Lines 173-181)

The use of a range of doses in phase 3 trials is put forward in this section as a recommendation to better characterize the relationship between exposure and the benefit-risk relationship, thereby "allowing provision of the best dosing advice." Studying a range of doses in phase 3 is reasonable if there are plans to market several doses for the disease. It is generally neither practical nor necessary for compounds that will be indicated for use at only one dose. In addition, the concept implies that current dose-ranging methods are somehow inadequate. Where the pharmacology and the PK/PD properties of a drug have been characterized, the value and, perhaps, the ethics of exposing large numbers of patients in phase 3 studies to doses higher than necessary or doses that are predictably less effective are questionable.

Somewhat contrary to the advice in this section, the concept paper appears to accept the value of PK assessments in determining dose-response relationships for both safety and efficacy in its discussion on detection of interactions (Section C, lines 212-214 - *"Including PK assessments allows for the determination of exposure-response relationships for both safety and efficacy...."*)

Recommendation: Current dose-ranging evaluation practices could be reviewed to identify their benefits and shortcomings - where they work well and where they are less useful. Recommendations for alternative approaches should focus on situations in which current practices are judged to be inadequate. Resources and ethics must be considered in making recommendations, particularly where vulnerable populations are involved.

6. III-C: How can unanticipated interactions be detected as a part of a safety assessment? (Lines 183-226)

a. Line 186 - "Clinical pharmacology studies do not guarantee a full understanding of all possible risks related to interactions."

No development program can achieve a "full understanding of all possible risks," nor is it a realistic goal.

Recommendation: In developing guidance, FDA should be particularly attentive to avoiding a "zero risk tolerance" posture.

b. For the reasons related to sample size mentioned in footnote 3, anything other than really large interaction effects are unlikely to be uncovered during later stages of premarketing development.

Recommendation: A more effective strategy would be to study potential interactions by specifically targeted Clinical Pharmacology or Phase 2 trials.

7. III-D: When would comparative safety data be useful? (Lines 228-251)

a. The paper indicates that one situation in which comparative studies (studies that include an arm with a well-characterized agent in addition to the test product) would be useful would be when there is a need to characterize background rates of certain adverse events. It is not clear how such a study design would be an improvement over placebo controlled studies to assess background rates of adverse events.

b. The section doesn't answer the question in the subheading. It describes when comparative safety data "could" be useful, not when it would be.

c. It is unlikely that comparative safety studies of any reasonably attainable size would be sufficiently powered to differentiate safety differences between available therapy and the new product.

d. There is an implication in lines 238-244 that when there is "well-established therapy," a new product must be safer or more effective than available therapy. As noted above, this is not a requirement of the FD&C Act. Furthermore, it is unlikely that any two drugs will have the same side effect profile varying only in intensity. It is generally accepted that, even within the same class, a choice of therapies is desirable. Some patients do better on one than on another.

Recommendation: In drafting the guidance document, FDA needs to present more clearly when good practice calls for comparative data in a development program with attention paid to the practical considerations such as study size and power to detect differences. It should also consider the regulatory application of such data, particularly when a single comparative study would suffice and when a replicate study would be necessary.

8. III-E. "What are some special considerations for optimal risk assessment during product development" - Lines 253-309

a. The "special considerations" section should be qualified to indicate that such measures are not necessarily requirements for initial approval but that they depend on the product and the disease under consideration.

b. Most of the strategies described will fail unless there is a fairly high risk of the event. Large simple safety studies may be helpful, but they would have to be very large to be effective.

Recommendation: A future guidance document must balance practical considerations to put risk assessment in the perspective of the overall drug development program, including other aspects of the risk management program.

9. III-F. "How can sponsors minimize medication errors?"

The concept paper recommends that sponsors perform a "medication error prevention analysis" or MEPA to identify known or potential errors and causes of errors and to minimize the potential for error by renaming, relabeling, or repackaging.

a. Lines 313-315: The concept paper lists the product's established name amongst the product attributes that, ideally, a sponsor would subject to risk assessment to assure against inadvertent contribution to medication errors. The established name serves a unique purpose in that is designed to convey information about the compound and to have similarity to other established names in the same class. This designed similarity is due to the "stem" system that requires the use of the same stem for products that are chemically or pharmacologically similar. Unlike the other items listed (proprietary name, labeling, and packaging), the established name is not the property of the sponsor but, instead, is assigned by the United States Adopted Names (USAN) Council, which includes an FDA representative in its membership.

The USAN Council cooperates with the World Health Organization (WHO) in working to establish a USAN that will be accepted by WHO as an International Nonproprietary Name (INN), the aim being to establish the same nonproprietary name worldwide. Both the USAN Council and WHO have principles for guidance in devising the established name. Any risk assessment would need to accommodate the necessary similarity in established names and the required principles of developing the established names set forth by the USAN Council and WHO, to which the sponsors must adhere.

The process of obtaining a USAN and INN is a lengthy one and the WHO group meets only twice a year. Any risk assessment would need to accommodate the schedule of both the USAN Council and WHO and the need for establishment of the name well before NDA submission.

While sponsors submit candidate names to the USAN Council, the Council is not obliged to accept the recommendations. The USAN process requires sponsors applying for a generic name to verify the absence of conflict with trademarks, other generics, and chemical names. Therefore, at Merck, established name candidates are subjected to a rigorous evaluation before applying to the USAN Council. The USAN Council further evaluates submitted name candidates for conflicts.

b. Lines 324-335 deal with recommendations for the evaluation of a product's name, labeling, and packaging. None of the recommended assessments are validated and, therefore, their

usefulness in contributing to reduction of medication errors is, at present, speculative. The value of these interventions would make a useful research program but it is premature to recommend them in a regulatory guidance.

Recommendations:

a. Because 1) the USAN Council requires sponsors to verify the absence of conflicts between established names they propose and existing trademarks, other generics, and chemical names; 2) the authority to adopt generic names in the United States is given to the USAN Council by the FDA, and the USAN Council coordinates with WHO, and 3) once assigned, the established name is neither owned by nor under the control of the sponsor, we believe it is inappropriate to imply that a sponsor is solely responsible for assessing risks associated with an established name. In any future guidance, we believe that recommendations for the assessment and prevention of medication errors associated with established names should be discussed separately from those pertaining to other product specific attributes in a manner that recognizes the existing USAN process and responsibilities.

b. Until methods for testing names and packaging have been validated to have predictive value in reducing medication error potential, it is premature to include recommendations for such testing in regulatory guidance.

c. Instead of recommending testing of unproven value, FDA should work with State medical boards and other State regulators of health professionals to address practice issues that are significant causes of medication errors. State laws defining the elements of a "complete prescription" to include specific directions for use would eliminate prescriptions with directions such as "as directed." A requirement to include the indication in the directions ("one tablet daily for blood pressure") would clarify many potentially misleading orders. Additional information in a prescription order is important to help identify the required product or to signal the need for clarification when handwriting or poor oral communication creates confusion.

d. FDA should direct resources towards evaluating and validating current and proposed testing practices to determine their legitimacy before seeking any recommendations for their use.

10. IV-C: How can analyses of dose effects contribute to risk assessment?

Lines 451-456 discuss products with recommended dosing that includes a stepped dosing algorithm such as incremental dosing based on age or weight. It suggests that it may be useful to make a specific effort to examine safety just above and below the cut points.

Recommendation: This type of analysis of cut points should be reserved for those products with a signal in the overall group, or a suggestion of a dose trend effect. Otherwise, the multiplicity of analyses will surely result in identification of a signal.

11. IV-D: What are appropriate methods for data pooling in risk assessment?

Lines 458 ff: It is unclear under what circumstances one might wish to carry out such an analysis at all. The integrated summary of safety (ISS) is normally where pooled analyses of safety are done. In addition, the requirement for a relatively homogeneous population in this section seems at odds with the argument for a heterogeneous database beginning on line 159.

12. IV-F: What is the role of subgroup analysis in safety assessment?

Lines 503-508 This section notes that demographic subgroup analyses are required by regulation and that other analyses may be of interest. While it acknowledges that most safety analysis is, to some degree, exploratory, it fails to emphasize the potential for identifying false signals.

Recommendation: To put subgroup analysis in proper perspective we suggest modification of the last sentence in this section to read as follows (added language in italics):

"They [subgroup analyses] have the potential to provide a more reliable and relevant estimate of risk for important subgroups of the target patient population; *alternatively, multiplicity issues could result in an apparent signal that may not exist."*

II. Concept Paper II. "Risk Management Programs"

A. General Comments

1. We commend the FDA for recognizing that the role of risk management planning is not to create a complex RMP for every product (Lines 122 - 124: "FDA anticipates that for most products that risk management planning will be handled by the information contained in the PI. Submissions to FDA to revise the PI for adverse events would not automatically lead to an RMP being proposed.") We encourage the Agency to emphasize this important concept in the guidance document.

2. Certain proposed risk management program (RMP) tools are, in effect, restrictions on prescribing and dispensing (physician/pharmacist certification or registration). By definition, a market application is approved by FDA if it is found that the information provided demonstrates the drug to be safe and effective "for use under the conditions prescribed, recommended, or suggested in the proposed labeling thereof." Approved prescription products may be dispensed by any licensed pharmacist upon receipt of a valid prescription from any practitioner licensed by law to administer such drug. There is no further requirement. Certain risk management plans would impose additional training and certifications on practitioners without which prescribing and dispensing certain products would not be allowed. This amounts to regulation of the practice of medicine and pharmacy.

3. Under certain proposed risk management programs, the lines between investigational drugs, approved products, and misbranded products may begin to blur. Under the FD&C Act, a product is misbranded if it is not safe and effective as defined in its labeling. Under section

201(m) of the FD&C Act, the term "labeling" means "all labels and other written, printed, or graphic matter (1) *upon* any article or any of its containers or wrappers, or (2) *accompanying* such article." (Emphasis added.) Continuing medical education, special training programs, and patient informed consent forms and other restrictive programs may not meet this definition. Indeed, defining a RMP as "a strategic safety program designed to decrease product risk by using one or more interventions or tools *beyond the package insert*" suggests that labeling alone is inadequate to assure safe use. Clearly, there are situations in which certain risks can be mitigated for a population for whom a product provides important benefits. In these cases, complex risk management programs may be justified by the need to permit access to such products by patients in whom the risk-benefit balance is favorable and for whom risks can be minimized. However, such risk management programs should be the exception, not the rule. Products subject to such programs must still meet the statutory requirements for approval.

4. The concept paper focuses on interventional tools that are part of an RMP and recognizes that for many, there are no reliable data that confirm their value. Accordingly, FDA is considering recommending pre-testing prior to implementation of an RMP. Given the long-term nature and objectives of risk management programs, useful methods for pre-testing remain undefined. Guidance should recognize the evolutionary nature of this field and avoid overly prescriptive recommendations that may be unattainable or of questionable value.

5. The other concept papers deal respectively with pre-market risk assessment and pharmacovigilance. There is a lack of cohesiveness between the three papers making it difficult to relate one to the other. We believe the Agency should consider issuance of a single, comprehensive document when it drafts guidance on this subject (see also our general comment on page 1).

6. The concept paper provides no discussion of how FDA will evaluate the overall public health benefit of an RMP and any unintended consequences thereof (possible effect of shifting prescribing to older, possibly less effective products that also have risks but are not the subject of an RMP). FDA's plans for assessment of RMPs should be described, including whether there are outcomes that would lead to stopping a program.

7. While the evaluation of the tools for RMPs is essential, employing best practices in the implementation of proven tools is also critical. For example, physician education is mentioned as a tool but how the education is implemented may affect the degree to which the additional knowledge leads to action in practice.

B. Specific Comments

1. III - When Would an RMP Beyond the Package Insert Be Appropriate?

Line 108: The document states that FDA may propose to the sponsor that an RMP merits consideration.

Recommendation: Given the implications of an RMP for the potential viability of a product and the resource requirements for such programs, we recommend that FDA promulgate a

reviewer guidance or MAPP providing general guidelines to promote consistency across Divisions and review teams for proposing an RMP and to create a level playing field for sponsors.

2. IV - What Interventions or Tools are Available for Use in Achieving RMP Goals and Objectives?

Line 178-180: "Instead of specific tools being presented as part of a guidance document, FDA may maintain a more easily updated resource on its website that describes tools that currently are in use."

a. The proposal to maintain a list of tools on the FDA website raises a number of questions: What would the criteria be for listing on the website? What does listing on the website imply? Does it imply acceptance of the tool by FDA or, indeed, a recommended approach compared to possible unlisted tools? Does it signify that the tool has been validated with respect to the objectives it would be implemented to attain? Would the posting on the website provide information to clarify these questions for each tool? Would conditions for implementation of the tool be provided so that it may be properly implemented?

b. The implication of an easily updated resource is that tools may be both added and removed from the list. What would the implication for ongoing programs be when a tool employed in those programs is removed from the list?

c. Use of the website instead of the guidance document to describe RMP tools raises the concern of implementation of substantive changes in guidance without notice and comment simply by changing the tools listed on the website.

Recommendation: The PDUFA III goals called for guidance on best practices. FDA guidance generally describes FDA's recommended approach to fulfilling regulatory obligations. We find it unlikely that the development of new, validated tools, that merit FDA's recommended approach, will progress at such a rapid rate that it will require constant and prompt updating. Furthermore, the specter of a constantly changing set of recommended tools is a concern given that a RMP represents a continuum throughout the development and marketing life of a product and, therefore, requires some stability and assurance that yesterday's recommended approach is not unacceptable tomorrow. For these reasons, we recommend that FDA describe the tools currently acceptable in the guidance. FDA might consider using the website to list new, promising tools that may be considered on a case by case basis. As experience with these tools leads to their validation, FDA may even post on the website information to show that those tools now have its full recommendation. The latter tools should be incorporated into the guidance at its next update.

3. IV-D. How does the choice of tools for an RMP lead to its broad categorization? (Lines 244 - 260)

a. FDA recommends that RMPs be "broadly categorized into one of several 'levels' to reflect how much the tools used in the RMP diverge from conventional prescribing, dispensing, and use." A four level categorization is described - Level 1, 2, 3, and 4. In the section that

describes the desired elements of a submission (lines 416-418), it is recommended that sponsors describe and categorize the overall level the RMP represents. There are no criteria proposed on which to determine the level into which certain types of risk might fit making the selection of levels entirely subjective while the imposition of an ordinal scale suggests standardization and comparability.

b. Even more critical is the way in which these categories will be used. These numbers, which over-emphasize risk over benefit, will be used by the press and the financial community and inappropriately used to categorize and stigmatize products. FDA has put forth no compelling justification for imposing a categorization system other than "for ready description and comparison of RMPs," an apparent bookkeeping function. Drugs will become "labeled" in the popular media as level 1, 2, 3 or 4.

Recommendation: We strongly oppose the proposal to categorize RMPs into a series of levels as described in the guidance. Such a scheme provides no additional benefits in terms of public health and safety. The predictable use of these numbers to inappropriately stigmatize products, as well as their focus entirely on risk, would be a source of disinformation for the public. It has the potential, in fact, to have negative public health consequences by steering patients and practitioners away from highly effective therapies based on misconception and confusion over the significance of the meaning of the level designation.

4. V - How and When Can Risk Management Programs Be Evaluated?

a. Lines 281-311: In section V(A) the document addresses pre-testing of programs. It also suggests that sponsors draw on previous experience from products with similar safety issues. While the concept of pre-testing RMPs is valid, it is far from clear how such testing might be implemented. RMPs are, in general, not short term commitments. This makes prospective evaluation of the probable effectiveness of such programs in achieving their objective very difficult.

Recommendation: Because FDA is the "clearinghouse" for all RMPs across all products, the Agency is in a better position than an individual sponsor to provide information on the usefulness of various RMPs based on experience, and thereby reduce the need for pre-testing. While sponsors with prior experience will, no doubt, rely on that experience in designing future programs, we recommend the Agency take primary responsibility for the collection and dissemination of information on the usefulness and limitations of RMPs, thereby reducing the need for pre-testing by every sponsor. In addition, we believe that the evaluation of risk management interventional tools is an area requiring further research and we suggest that FDA consider entering into a collaborative effort with organizations such as CERTS to investigate evaluation methodologies.

For products where a program is being designed during development and where pre-testing is considered necessary, FDA should provide additional guidance on the kind of pre-testing expected, if any, beyond comprehension and, perhaps, willingness of practitioners to participate. Such guidance should include clear objectives that are both practical and attainable.

b. Lines 343-384: Sections C (How can RMP effectiveness be measured?) and D (What are the strengths and limitations of different evaluation methods?)

The issue of evaluation of the RMP raises a number of important questions that are not addressed in the concept paper. These include how one defines success and the threshold of acceptability. In addition, the timeframe for evaluation of the program needs consideration. How soon should evaluation be attempted? How often? And for how long?

Evaluation of risk reduction programs as described in the concept paper is not a trivial task. If a sponsor has, through experience, put in place a set of 'good practices' for minimizing risk, it would seem unlikely that further changes in the risk prevention strategy would produce material decreases in risk.

If the improvement is large enough, it should not be difficult to demonstrate the benefit of the new practices. However, following introduction of effective new practices, it will likely be much more difficult to demonstrate that any changes in procedure will pay off in less risk or, more importantly, that any changes that might be observed are not due to chance alone. Identifying, and verifying, strategies that provide marginal risk reduction will be very difficult when there is not much room for improvement and when the practical evaluation measures are confounded and imprecise.

The concept of risk management programs and evaluation thereof is an evolving field with which there has been limited experience to date. The recommendations for conducting RMP evaluation seem speculative and lacking in practical, achievable results, particularly when considering a prospective risk management under which a product would be marketed from the day of its approval. In such a situation, there would be no baseline event rate against which to measure the effect of the plan.

Recommendation: It may be more productive to identify through guidance the components of good practice in product usage, some of which are in the concept paper, and simply implement them in selected high-risk therapeutic areas. It is particularly important not to allow the deadline in the PDUFA III goals for publication of guidance on this subject to lead to hasty recommendations. The guidance should represent current thinking on best practices. Where best practices have yet to emerge from experience it should be acceptable for the guidance to so specify.

III. Concept Paper III. "Risk Assessment of Observational Data: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment"

A. General Comments

1. It is not clear that a careful assessment of current pharmacovigilance practices to identify their advantages and shortcomings has been undertaken. Without such an assessment as a first step, the task of identifying best practices cannot be properly accomplished.

2. Hypothetical risk management solutions and practices that have not been validated should not be part of the future guidance document.

B. Specific Comments

IV: How are Safety Signals Best Evaluated in Pharmacoepidemiologic Studies?

Lines 135-137: It is stated that "When a safety signal is identified from spontaneous case reports, literature reports, or other sources, further evaluation of the signal may be possible via carefully designed pharmacoepidemiologic studies, registries, or surveys."

Recommendation: In addition to the methods mentioned, it should be noted that a review of preclinical data as well as controlled clinical trial data from well designed clinical trials may also be useful for further evaluation of the signal, as well as to balance the perspective of a spontaneous report.

Line 182: According to the concept paper, validation of diagnostic findings in claims database studies through detailed review of at least a sample of medical records is essential for all pharmacoepidemiologic studies. However, there may be circumstances in which use of pharmacoepidemiologic databases might provide valuable information without having to go back to the medical record.

Lines 211-235: This section addresses when and why surveys would be performed. Included in the list of applications for surveys is to "address confusion in the practicing community over sound-alike or look-alike proprietary names." It is difficult to understand how a survey could "address" such confusion, as in mitigating or ameliorating it. While, theoretically, surveys might be used to "assess" such confusion, such attempts have never been validated and, therefore, are unreliable.

Recommendation: It is essential, in drafting guidance documents on risk assessment and management, that FDA avoid endorsing any tools or practices for which the value has not been demonstrated. Such recommendations may result in inappropriate use of limited resources and unnecessary interventions that are ineffective.

Lines 288-296: The opening sentence of this paragraphs states, "To provide context for incidence or reporting rates, it is helpful to have an estimate of the background rate for the event being evaluated...."

Recommendation: Incidence rates should not be grouped with reporting rates. Comparison of incidence rates to background rates is far less problematic than the comparison of reporting rates to background rates.

Lines 295-296 state that "...a high reporting rate may, in some cases, be a strong indicator that the true incidence rate is sufficiently high to be of concern."

Recommendation: If similar language is included in the guidance document, other reasons for a high reporting rate should also be noted to make it clear that a high reporting rate alone

may also be generated by other factors unrelated to a high true incidence. For example, a high reporting rate may be a reflection of media attention, the newness of a product in the market, and other reasons.

2. V: How Are Safety Signals Best Assessed and Interpreted

Line 347-349: "FDA has found that it may be possible to assess the degree of causality between use of a product and an adverse event when the sponsor gathers and evaluates together all available safety data...." We agree with this statement. Since adverse experiences become apparent in 3-6 months post-approval, access to all available data is important to a Company.

Recommendation: FDA needs to review the process whereby direct reports made to FDA for all products are readily available to companies to aid them in understanding the safety profile of their products.

Lines 358-369 recommend follow-up with reporters of medication errors associated with a safety signal to determine the "root cause factors" that led to the event. Lines 364-369 describe the characteristics of this analysis, including identifying failure points in the medication use system, looking beneath "visible cause", and identifying prevention strategies.

Prescribing and other process errors that are unrelated to a specific product characteristic and, therefore, out of the control of the manufacturer should not require a comprehensive root cause analysis by the manufacturer of the product involved. This recommendation extends FDA's regulatory authority over products to a broader regulation of health care practice. Pharmaceutical manufacturers have no authority to access the information necessary to perform a root cause analysis of private medication use systems.

Recommendation: The responsibility for identifying "failure points" in the systems of health care institutions or private providers does not rest with the manufacturer of the drug involved and, indeed, the product manufacturer does not have access to information about these private systems. Medication use system failures are medical, pharmacy, and health system practice issues. Investigation of these systems and their regulation rests with State licensing authorities. This recommendation should not be incorporated into FDA guidance on good pharmacovigilance practices and pharmacoepidemiologic assessment.

Lines 371-372: The concept paper recommends the use of the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Taxonomy for classifying and tracking errors.

Recommendation: Agency guidance should not incorporate the standards of outside organizations into guidance. This has the effect of transferring regulatory authority to the organization because the organization may change its recommendations at any time, thus changing the guidance.

Line 376-381: Section V(D) entitled, "How would safety signals be reported to FDA?" describes a process that, on its face, would appear to significantly delay updating the labeling

of products compared to current practice. As described, when a safety signal is identified, sponsors would be expected to submit a synthesis of all available safety information (pre-clinical to current observation), provide a re-assessment of risk-benefit of the product, propose further investigation, and propose risk management strategies. Upon receipt, FDA would make its own assessment. In contrast, under the current system, when a potential safety signal is identified, assessed, and determined to warrant a labeling change, a "changes being effected" (CBE) supplement is submitted and the labeling is revised right away.

Recommendation:
a. Rather than require a full synthesis of every potential safety signal, a more reasonable approach involving dialogue with the Office of Drug Safety and the Review Division should be considered to discuss plans for further investigation.

b. Guidance should include in this discussion some reference to the timing of labeling submission in the event that the sponsor has made the decision, on the basis of its evaluation of the signal, that a labeling revision is called for.

3. VI: How Can Safety Signals Be Monitored Through Enhanced Pharmacovigilance Efforts?

Line 452: In the following paragraph FDA seems to be using the term "signal" instead of the term "risk" (italicized below):

"While additional information is being developed, FDA may decide to take interim regulatory action to communicate information about safety signals via labeling or other means, or minimize the *signal* in users of the product via risk management strategies as discussed in the concept paper entitled *Risk Management Programs.*

Recommendation: Assessment of a signal is essential before considering regulatory action. Without proper understanding of the signal (based on scientific evidence) an interim regulatory action has the potential to create unwarranted concern in the public with potential public health consequences. Because of the potential for unintended consequences of premature action, it would be useful if regulatory guidance provided further discussion of the kinds of situations in which such action may be considered. In addition, we encourage inclusion of a recommendation for dialogue with the sponsor when regulatory action is being considered prior to completion of a full assessment of a signal.

C. Answers to FDA's Specific Questions

How can the quality of spontaneously reported case reports be improved?

ANS: Education; identification of best and worst practices; simplify reporting.

a. There are two aspects to spontaneous reports - the technical aspect (assuring that all elements are present) and the clinical information (the information necessary to allow evaluation of the report). The provision of adequate clinical information is a shared responsibility with practitioners. There is a need to increase awareness by healthcare

professionals of the importance of pharmacovigilance. To this end, there should be education of practitioners to communicate high quality data with specific clinical details, and education of patients to recognize signals and report specific details.

FDA should continue to partner with other organizations, including PhRMA, and with medical, pharmacy, and nursing schools to increase awareness of the importance of pharmacovigilance and clinical pharmacology. Programs should emphasize the importance of prompt reporting and the basic elements necessary for a complete and evaluable report. Further partnership with health professional organizations should also be forged to reinforce these principles. Continued collaboration with the National Patient Safety Foundation to improve patient education on complete and accurate ADR reporting is also important.

b. Analysis of what works and what doesn't: FDA should provide perspective on the aspects of case reports, their sources, and type of follow-up they have found most helpful in signal detection as well as an analysis to identify interventional tools in use that are not productive.

c. Simplify reporting: Safety reporting may be viewed by busy practitioners as burdensome and time consuming. The minimum information necessary to provide adequate information to evaluate a report should be defined with the goal of making the reporting process as simple as possible.

2. What are the possible advantages and disadvantages of applying data mining techniques to spontaneous report databases for the purpose of identifying safety signals?

ANS: Advantages: Data mining techniques provide a systematic, efficient approach to screening large adverse event databases for potential safety signals, particularly involving rare events, drug interactions, and demographic factors. They improve the ability to evoke the emergence of associations between events and concomitant variables.

Disadvantages: The limits of the underlying data and the data mining techniques must be fully appreciated to avoid false positive causality conclusions which could lead to costly interventions and inappropriate product restrictions.

3. What are the possible advantages or disadvantages of performing causality assessments at the individual case level?

ANS: Advantages: With the exception of cases that include a positive re-challenge, there is little or no advantage in performing causality assessment on individual case reports.

Disadvantages: High likelihood of misinterpretation

Given the inadequacy of much spontaneous report data, application of causality algorithms to a simple case is fraught with misinterpretation. The ability to rule out the likelihood that the suspect drug may have contributed to the adverse experience, in most instances, becomes impossible. Therefore, most adverse experiences at the individual case report level end up with a possible association. The one exception to this is the case involving a positive re-challenge.

4. Under what circumstances would a registry be useful as a surveillance tool, and when would it cease to be useful?

ANS: Registries may be useful when (1) the information you need to assure safe use of the product cannot be obtained from any other mechanism, and (2) the need for the information justifies the extensive resource commitment necessary to successfully set up, operate, and comply with the registry.

A registry allows determination of a denominator which is potentially helpful when the outcome of concern is relatively uncommon and there is an agreed upon background rate that is applicable to the population captured in the registry. There should be a time frame or stipulated number of exposed patients so that the effort is not an open ended one that never leads to a conclusion. In addition, one needs to think about the settings and types of registries as some already established registries, such as cancer registries, may not capture the exposure or outcomes data of interest.

It is neither necessary nor appropriate to create a registry for most newly approved products. Registries need a clear focus and specific objectives. One needs to balance access to the drug with level of concern. Alternatives to registries should always be considered. A well-tested alternative to registries is the use of large automated databases to serve as a source of appropriate cohorts. Databases such as Kaiser Permanente, Saskatchewan in North America, and General Practice Research Database (GPRD) in Europe are powerful tools used by industry to sponsor research without assuming the cost of initial data collection which registries entail.

5. Under what circumstances would active surveillance strategies prove useful to identify as yet unreported events?

ANS: Active surveillance strategies, which often include the concept of registries, are mechanisms for getting information you cannot obtain from any other tool. They should be reserved for situations in which the information needed cannot be obtained from other, less resource intensive and less obtrusive tools.

Active surveillance strategies that involve use of sentinel sites may be subject to selection bias and may involve small populations. When successfully employed, however, they may provide more in-depth information, a more accurate estimate of the reporting rate, and may be more likely to provide information that will lead to a causality assessment.

6. Under what circumstances would additional pharmacoepidemiological studies be useful?

Pharmacoepidemiologic studies are useful to further define hypotheses generated by a signal from the spontaneous reporting system or to further evaluate the safety profile of a product for which a potential risk has been identified prior to approval. They are also useful to examine background incidence rates of risk factors for the AE of interest identified in the spontaneous reporting system.

Drug utilization studies are useful to study physician prescribing behavior, drug use and impact on outcome, and to evaluate risk management strategies for unintended consequences. They may also prove useful to evaluate the effectiveness of Risk Management interventions.

Conclusion

We commend the Food and Drug Administration for issuance of the three concept papers on risk assessment, risk management, and good pharmacovigilance practices and pharmacoepidemiologic assessment. These documents, along with the public workshop on April 9, 10, and 11, represent an extraordinary effort on the part of the Agency to convey its preliminary thoughts on these issues and to stimulate discussion with stakeholders.

We recognize that the concept papers are, themselves, not intended to be guidances on best practices in risk assessment and risk management and that, therefore, it is appropriate to explore, in these papers, a broad range of tools, including those with which there has been little experience. In preparing actual guidance on best practices, however, it will be important to narrow the focus to practices, the value of which has been documented by prior experience, and for which their role, their advantages, and their limitations can be articulated.

The call for guidance on risk assessment, risk management, and pharmacovigilance activities in the PDUFA III goals is neither an expression of concern that current efforts are inadequate nor a call for more intense surveillance. It simply a call to document those practices that represent the best of what we are doing now. Risk management, itself, is not new to drug development. As an industry, in conjunction with the FDA, we have been conducting pre-approval tests of increasing intensity and complexity on potential products for decades; we have been collecting, monitoring, and evaluating spontaneous reports on marketed products and taking appropriate action to minimize risks. Likewise, we have carried out phase 4 programs based on commitments made to the Agency at the time of approval to address potential, often theoretical, risks that had not been resolved at the time of approval. It is the best of these practices that the guidance is intended to capture.

The field of risk management is evolving, however, and the goals of the risk management programs of tomorrow may be different from the goals of the past. New ideas on how those goals might be reached constitute an area of risk management that requires research to define and evaluate new programs. It is important that these new horizons be explored, both by industry and by FDA, and perhaps in conjunction with organizations such as CERTS, but new ideas, no matter how appealing they may appear, should not be confused with best practices and should not be incorporated into the guidance called for under the PDUFA III goals as though they were tried and true techniques.

We welcome the opportunity to comment on the development of this important guidance. We have provided both detailed comments on concepts articulated in the three concept papers as well as our general thoughts for the future guidance. We believe that a single final guidance combining risk assessment, risk management, and good pharmacovigilance and pharmacoepidemiologic assessment practices would provide a more coherent approach than three separate documents.

We trust you will find our perspective useful in your efforts to draft guidance on best practices in risk assessment and risk management. If appropriate, we welcome the opportunity to meet with you to discuss these issues further.

Sincerely,

David W. Blois, Ph.D.
Senior Vice President
Global Regulatory Policy